---------- Forwarded message ---------
From: **John Fuerst** <j122177@hotmail.com>
Date: Mon, Jul 20, 2020 at 10:32 AM
Subject: Re: Intelligence study
To: Russell Warne <russwarne@gmail.com>

Russell,

Bryan and I resubmitted the TCP Hispanic paper the other day.

As agreed, we left your name on. However, in the unlikely event that it will be accepted, we will take your name completely off as agreed.

Good luck with the new kid.

Best,

John

From: Russell Warne <russwarne@gmail.com>
Sent: Saturday, July 11, 2020 4:33 AM
To: John Fuerst <j122177@hotmail.com>
Subject: Re: Intelligence study
John,

The fact that I haven't been able to even look at my personal email account for two weeks shows how little time I have right now. Because of the pandemic, my wife and I haven't had as much help as we did when our other children

were born. Plus, there are now more kids than ever in the house. My biggest priority right now is taking care of my family. I only get about 3-5 hours per day to work, and that means that I am extremely delayed in everything. (When you work part time, nothing is on time.)

I just can't commit to any project right now--or for the next several months--and need to be taken off it completely. It's not fair to leave you dangling like this.

-Russell Warne

On Thu, Jun 25, 2020 at 12:16 AM John Fuerst <j122177@hotmail.com> wrote:

> Dear Russell,
>
> Thank you so much for the update regarding your schedule. As agreed, we will take your name off during a later round if you end up not contributing substantially to the review reply.
>
> As noted, Emil and Davide are uninterested in helping. So that leaves Bryan and potentially you. I have done what I could. But it needs another hand or two to work on it.
>
> I don't know what Bryan's time frame is. As it is, we are doing a resub for another paper & working on other stuff.
>
> We will keep you updated on the situation with that paper. I don't expect it to get peer approved this year, though.

From: Russell Warne <russwarne@gmail.com>
Sent: Thursday, June 25, 2020 5:01 AM
To: John Fuerst <j122177@hotmail.com>
Cc: Bryan Pesta <bpesta22@cs.com>
Subject: Re: Intelligence study

John,

Because of my new baby and the demands of my job, I am not going to be able to help with this revision for the foreseeable future.

Please take my name off the manuscript. It's not fair for me to leave you hanging like this for a few weeks, especially after saying that I would need to be a more involved party in the manuscript. Given my current time crunch, I am unable to say when I would be able to give the manuscript the attention it would need.

-Russell Warne

On Sat, Jun 20, 2020 at 11:18 PM John Fuerst <j122177@hotmail.com> wrote:

> Ok Guys,

I have to update table numbers and stuff, but I rewrote the paper.

I need you to read it over in conjunction with the tentative review replies.

Paper:

https://docs.google.com/document/d/1pO_hXFY-Tm1H7pvOxxMbvJiXxs7GDfk1VgRlgggj_CU/edit?usp=sharing

Replies:

https://docs.google.com/document/d/1_coMd9-lJHiYDEfYHTcE9XnIDpcGYijraS_eDFr_ixQ/edit?usp=sharing

Give me some feedback and then we can hammer out what we want to say.

Based on our recent UV paper rejection (the 3 original reviewers agreed, but a hostile 4th was brought in in the second round who was deeply concerned), I would give this a 10% chance. Maybe lower given the current political climate. But I am going to try.

From: John Fuerst <j122177@hotmail.com>
Sent: Friday, June 5, 2020 9:25 PM
To: Russell Warne <russwarne@gmail.com>
Cc: Bryan Pesta <bpesta22@cs.com>
Subject: Re: Intelligence study

Dear Russell,

If I did not send you the replies, that is my fault. Professor Haier gave us a chance to reply. But the reviewers asked for a lot and, as he noted, it may be impossible to please them.

Emil, Davide, Bryan, and I were split on whether we should bother. However, knowing how difficult it is to find a journal that will not editorially reject this type of research, I decided to go for it, but since that

was my decision I am getting minimal help on this from Emil and Davide.

So, I have begun the long process of revision:

https://docs.google.com/document/d/1pO_hXFY-Tm1H7pvOxxMbvJiXxs7GDfk1VgRlgggj_CU/edit?usp=sharing

(See the link on the top for my replies in progress.)

I would suggest to give me a couple of weeks to outline a strategy. (The response has to be done in a clever way, though, since we lost our imputed data files and are limited in what additional analyses we can do.)

And then, when you have a chance in a couple of weeks, to look it over. If you feel you can materially help and do actually help we can leave you on. And if not, we can keep the default agreement of removing you.

There are a number of statistical issues raised, like the effect of unbalanced samples, which you may have an idea about. My understanding of them is, frankly, weak -- and Emil, who I usually ask, is averse to

helping, since he voted against resubmitting.

John

.....

"Richard Haier

Editor

Intelligence

Editor and Reviewer comments:

Reviewer #1: The paper is interesting and ambitious, conducting a range of analyses aimed to test hypotheses via controlling for potential biases and covariates. The analytic plan is appropriate, however 2 major limitations preclude from making solid conclusions or ability to disentangle some potential explanations: limited sample (with a wide age range) and absence of a more detailed SES measure. I outline my concerns and comments below.

Main concerns:

- Descriptive statistics for cognitive measures for all sub-samples are required. Much of the discussion and interpretation depends on these measures. It is important to examine their validity for different ages, how g measure was constructed, the observed differences etc.
- Overall, the results are limited by the fact that SES was only measured by parental education, and therefore socio-economic variation (including disadvantage) was not tagged. The authors mention in the

discussion that associations with other unmeasured variables may explain the links between g, ancestry and PGS and that this must be explored further. This is a very important point as indeed ancestry is linked with privilege and socio-economic advantage, which in turn is linked with g. The conclusions of the present study are extremely limited without a more extensive measure of SES.

Specific comments

1. Abbreviation MTAG appears without introduction in the abstract.
2. The sentence about attenuation of effect ...'not further after adding phenotype' is not clear: what phenotype? This also needs to be clarified in the Highlights and the Title.
3. Explain why only Hispanic sample's results are mentioned in the Abstract. Overall, the title can better reflect the focus of the paper.
4. Introduction. The following sentence requires revision: 'One explanation appeals to putative, lower within-group heritability in non-White groups, since predictive accuracy is a function of this variable'
5. The following statement requires supporting sources: 'and because sub-Saharan Africans are the continental lineage most genetically distant from Europeans (and other major races).
6.The following point requires further explanation: 'This is because lower p-value SNPs are theoretically more likely to be causal (Spencer, Cox, & Walters, 2014; Wang & Teo, 2015; Grinde et al., 2018) because causal SNPs are more transethnically valid when common variants are involved (Marigorta & Navarro, 2013).
7. The paragraph on page 3 starting: 'Among admixed populations...' is somewhat difficult to follow, the flow must be improved.
8. 'Since the relation between ancestry and eduPGS is better characterized as constitutive, rather than causal...' Please explain this further.
9. Clarify this: 'as would be expected were genetic g largely underwriting the association among both African and European Americans'
10. Methods: wide age ranges in both samples, but the samples are of very different size. The methods should describe suitability of the chosen instruments for children and adults; discuss the age range in each explored sub-sample.
11. From the method section, it is not clear how g was calculated from the available subtests. More discussion is needed about observed group differences for different subtests. There needs to be another table with descriptive statistics for each cognitive measure.
12. The regression analyses are limited as different models include different numbers of participants. The authors discuss the lack of association with lack of power, but this is speculative. Also, the effect from SES, this needs to be discussed.
13. The authors are correct in this statement:' However, the total Amerindian ancestry variance involved here is very small, so not much can be made of these results. Generally, it would be worthwhile to attempt replication...' This also applies to other results as the two samples and (subsamples) were not comparable in numbers.

14. The regressions for PGS on subsamples are really limited in terms of Ns and are unlikely to be informative.

15. In the path diagram the positive association between skin colour and g was of similar magnitude to the PGS association with g and some other associations. This needs to be discussed further. Otherwise, the authors run the risk of selectively either under-interpreting or over-interpreting similar size associations.

16. 'As adolescent European ancestry does not cause parental education, the relation is represented as a covariance (indicated by dashed lines).' This logic is not clear. Ancestry is not adolescent, and indeed can contribute to education of parents.

17. 'Spearman's hypothesis' and 'Jensen Effect' need to be expanded on. They are not adequately introduced in the introduction, and not clear in the discussion.

18. The attempt to control for bias by excluding variants < .05 MAF, may not be stringent enough in such a small sample.

Reviewer #2: I have decided to recommend major revisions of this manuscript. I doubt that the authors will be willing to undertake these revisions, but they should be given the opportunity.

Briefly, this paper is arguing that the EA polygenic score (PGS) has predictive power in Hispanics and that the European-Hispanic difference in IQ is explained to some extent by the difference between these populations in the PGS. Perhaps the paper does not state the latter point in so many words, but that is what it amounts to. The major problem with respect to this implication is that it does not address the problem of population stratification -- confounding in the original GWAS, such that alleles with higher frequencies in Europe tend to be spuriously associated with the trait. This problem is particularly severe in the GIANT Consortium's GWAS of height, which when used uncritically in population comparisons leads to the dubious prediction that Europeans should be much taller than Africans (Martin et al., 2017; Berg et al., 2017). I think it is possible to address this problem. There are two recent papers that have thoroughly examined the stratification affecting the
GIANT results (Sohail et al., 2019; Berg et al., 2019). One might perform each of the analyses in these papers, except applying them to the EA results rather than the GIANT Consortium's height results, and showing that stratification is present to an immensely smaller degree in the EA GWAS. (Some of the analyses require looking only at the UK Biobank. This is possible with educational attainment. Other analyses require looking only at within-family results; this is possible if the EA3 authors are willing to cooperate, or if the authors of the manuscript under review are willing to wait for the fruition of ongoing research.) One might have to creatively extend these analyses to deal with the African and Amerindian populations from which admixed Hispanics derive their non-European ancestry. But if one could show that there is no general genome-wide spurious tendency for alleles at higher frequency in Europeans, relative to the other source populations, to be associated with
higher values of cognitive phenotypes, then one could convincingly argue that European-Hispanic gap in IQ is indeed genetically caused.

Here are my comments tied to specific parts of the manuscript.

abstract: "trybrid" This is an unusual word. Why not just "admixed."

pp. 2-3: "This is because lower p-value SNPs are theoretically more likely to be causal (Spencer, Cox, & Walters, 2014; Wang & Teo, 2015; Grinde et al., 2018) because causal SNPs are more likely to be transethnically valid when common variants are involved (Marigorta & Navarro, 2013)." --> "This may be because lower p-value SNPs may have certain properties that increase transethnic validity, such as being common SNPs likely to have been segregating in the ancestral human population (Marigorta & Navarro, 2013)." This original sentence is clearly malformed. Also, some readers might be confused because they may be aware that a SNP's LD with neighbors is predictive of its GWAS association strength, which might naively suggest that highly significant SNPs simply tag more SNPs and are less likely to be causal themselves. My suggested rewording does not give that impression and also more accurately captures the content of Marigorta and Navarro (2013).

p. 3: "trybrid American (European-African-Amerindian)" --> "Latin Americans with European, African, and Amerindian ancestry" I suggest avoiding unusual words or neologisms.

p. 16: "As seen, the line deviates from linearity at six points between 50% and 75% European ancestry." What does this mean? I suggest leaving it out.

p. 18: "As a result, the bivariate correlations between ancestry and outcomes are attenuated via restriction of range." Here and elsewhere, why use correlation when this is affected by the variance of one variable? When there is a natural choice as to which variable should be X (e.g., ancestry), why not use the regression coefficient, which is not affected by restriction in the variance of X?

Tables 7 and 8: It is convenient to put everything in a correlation matrix. But the associations between PGS and cognitive ability are more interpretable when expressed as regression coefficients. As stated earlier in the paper itself, the correlations might differ between groups simply because the allele frequencies differ.

p. 25: Figure 6 alleviates my previous concern somewhat. However, is the difference in intercept between the group statistically significant? To my eye, they may not be.

p. 28: I am confused by this path model. How can one manipulate the variable EUR? That is, how can I make someone more EUR, and thereby affect their PGS, skin color, and g? I mean conceptually, not practically in a real experiment. The authors earlier cite some philosophy papers about constitutive rather than causal relevance, and I think I have an inkling of what this means, but this might not be enough for some readers. I suggest leaving out this analysis, which does not seem absolutely critical to the paper. For example, simply regressing IQ on the PGS and skin color is equivalent to a path model where PGS and skin color send arrows

to IQ, and PGS and skin color are correlated for unspecified reasons. And even if the authors stick to path modeling, some readers might wonder why other path models are untried. How about the causal chain PGS --> color --> g, as well as the direct path PGS --> g?

p. 32: I find some of the correlations in Table 12 too high to be believable. The correlations of the PGS with subtests surely are affected greatly by sampling variation. And yet the g loadings show a correlation with PGS r of 0.90 even in the Hispanics? Can the authors do something to reassure readers that there is nothing fishy here? Perhaps a scatter plot with g loading on the x axis and the other vector on the y axis and 95% confidence intervals in both directions tacked on to each data point.

p. 34: "To investigate, we compute eduPGS by derived and ancestral status." It is great to see that the authors deal with the potential bias noted by Kim et al. (2018). But the authors leave out some important details. Do the authors compute one PGS with only those SNPs where the enhancing allele is derived and then another PGS with only those SNPs where the enhancing allele is ancestral? Kim et al. (2018) speak only of "risk alleles" and so it is hard to tell just from the earlier paper what exactly the authors did themselves. Also, the authors' formula for weighting these two PGS does not obviously follow from Kim et al. (2018, p. 12). Are the two methods equivalent? If so, can this be explained more clearly? Also, perhaps pass along the caveat of Kim et al. that their method might not correct all forms of allele-frequency bias.

pp. 35-38: These analyses is far too sketchily described. How were the population-specific betas obtained? The authors must have performed a GWAS of European and African Americans separately, presumably in the TCP. Where are the details of this? What were the sample sizes? What was the statistical power to obtain a concordant beta? Etc. Even if statistical power is enough to get more out of this method than noise, how do we know that the better validity of the concordant PGS is not driven largely by the fact that its SNPs tend to have higher minor allele frequency? If that were the case, it would be useful to know this. More importantly, many readers are going to find it hard to believe that weighting the SNPs by effect sizes ascertained in small-sample GWAS is going to get you anything. The authors will have to do power calculations to assure such readers that the results are reasonable. Also, the authors will need to explain why the racial gaps increase when the
ethnic-specific betas are used.

p. 35: "neither of these three common forms of ascertainment bias could" --> "none of these three suggested forms of ascertainment bias could"

Berg, J.J., Zhang, X., & Coop, G. (2017). Polygenic adaptation has impacted multiple anthropometric traits. bioRxiv. doi:10.1101/167551v4

Berg, J.J. et al. (2019). Reduced signal for polygenic adaptation of height in UK Biobank. eLife, 8, e39725.

Sohail, M. et al. (2019). Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. eLife, 8, e39702."

From: Russell Warne <russwarne@gmail.com>
Sent: Friday, June 5, 2020 8:50 PM
To: John Fuerst <j122177@hotmail.com>
Cc: Bryan Pesta <bpesta22@cs.com>
Subject: Re: Intelligence study

John,

I will consider your point when the reviews come back. For now, the default decision is to remove me from authorship. But my mind can be changed if (1) I do the work, (2) the changes needed are within my expertise to do, and (3) I feel comfortable doing so. Who knows what the situation will be like in the next few weeks?

-Russell Warne

On Fri, Jun 5, 2020 at 2:34 PM John Fuerst <j122177@hotmail.com> wrote:

> Dear Russell,
>
> We agree to remove you from the paper, unless you change your mind, but at a more expedient point in the process.
>
> That said, if the concern is genuinely or simply that you feel your authorship is not presently justified, why not contribute with the difficult review reply and the overhaul of the introduction and discussion! The latter have to be rewritten in a way to convey what I intended but apparently did not get across: that results only showed that eduPGS, ancestry, and g are statistically tangled, which could be due to either cause or confounding & that the cause vs. confounding issue needs to be resolved, in future research, to get accurate eduPGS estimates.
>
> Unfortunately, neither Emil and Davide are interested in making these nuanced arguments. And Bryan is a boomer, not really familiar with the research. He has trouble even plugging in his printer. And I am having trouble with this due to brain fog and headaches. Now, I have a feeling you will say, "Time!" But critical commentary and help with this will not take so much of that. And yet given the tenuous status of this

paper, and the heavy criticism the reviewers are subjecting it to, it is so necessary. Frankly, I could use your help.

To be clear, with the above, I want to remove from your mind the primary reason you are giving and force you to make a difficult decision: either help with this research which you yourself advocated for because of its scientific and social importance *or* abdicate some of your responsibility as a knowledge-seeker because of possible social repercussions.

I do see this research topic as a sort of moral test for academics, almost all of whom have failed.

Anyways, mull over my suggestion. It will take me a few more weeks to hammer out a very rough draft, since I am going slow for the previously mentioned reasons. So if you change your mind, let me know by the end of the month and I will send what I have for commentary/review / help. Otherwise, I will just go with the best I can do, inadequate as that may be.

John

From: Russell Warne <russwarne@gmail.com>
Sent: Friday, June 5, 2020 6:56 PM
To: John Fuerst <j122177@hotmail.com>
Cc: Bryan Pesta <bpesta22@cs.com>
Subject: Re: Intelligence study

John,

I don't care when my name is removed. If it makes it easier to remove it after the pending decision from Intelligence, then so be it.

While I was part of the early talks and I did get dbgap access, I did not do anything else. Even if the controversy weren't enough to give me cold feet, it still doesn't change the fact that I do not feel that I have contributed enough to warrant authorship. I agreed to authorship a few months ago for short-sighted reasons (wanting to fatten the vita, mostly), and that was wrong. But even before the riots, I had been ethically uneasy about my coauthorship. I'm not one of those scientists who has his name slapped on a hundred articles that he barely had a hand in. For every article I've coauthored, my contributions have been so important that the article would not have been written without my efforts. I should not have deviated from that track record in an effort to fatten my vita.

I'm sorry if this inconveniences you or makes things more difficult for you.

-Russell Warne

On Thu, Jun 4, 2020 at 8:51 PM John Fuerst <j122177@hotmail.com> wrote:

Dear Russell,

Bryan forwarded your email and requested I respond.

Prior, let me apologize upfront for being frank to the point of rude -- I have been having bad headaches again and these make me irritable.

The short answer is, "no," at least not at this stage of review. However, if it gets accepted at *Intelligence*, which is unlikely, we will notify Professor Haier that you have been taking flack and would like your name taken off of the manuscript, if you insist. And if, in the much more likely circumstance that the paper gets rejected, we will not include you when submitting to another journal, again, if you insist.

My reason is practical. Both you and Bryan were identifiable to reviewers in the paper by virtue of the statement of data access. I don't wish them to suspect something amiss as they might if we changed that and also deleted the IRB approval sentence. Moreover, I am of the opinion that Professor Haier gave us a chance to reply in part because we had as a co-author an *Intelligence* board member. I don't wish to give him reason to change his mind. I spent and incredible amount of time and effort and a substantial amount of money on this. I am not going to risk this slight opportunity over your qualms.

As with you, I believe we are in the right to do this because you agreed to be a coauthor on this paper numerous times over the past year. Moreover, there was no violation of research ethics in either letter or spirt, in context to this specific paper, as it is not true that you had "no hand in the creation, execution, or write-up of the study" or that you were "a front to gain access to the data". So that is not a sound excuse.

As to the former point, if you check your email history you will see that about a year ago I had encouraged you to apply for dbgap TCP data. Then, after Bryan and I had published Lasker et al. (2019), you wrote, congratulating us, and saying that we had beat you to the analysis. At that time I raised the possibility of a paper on TCP Hispanics. We agreed to collaborate. This was even before I raised the idea with Bryan. Therefore, this project was your and my original conception. Thus, you had an undeniable hand in its creation. Though, if you feel guilty about leaving the hard work to me with my disorder of mitochondrial metabolism and consequent chronic lack of energy, I will be happy to send you the many page reply to reviewers, in a few weeks, for you to check over so to help alleviate both my burden and your conscience.

As for the latter point, you were not, since you never gave us access to data. Regarding data, the agreement was that *if* reviewers requested novel analyses which Bryan I and could not do then we would somehow go through you as a backup. Yes, this somehow could have potentially involved you

being a "front", as opposed to me sending you the code to run. But this is a moot issue. Our imputed datafiles were lost: a complete hard drive crash. So, given time constraints, I can only try to address the reviewers requests using 1000 Genomes data and the saved output (computed variable) files which we have. However, since I have only this to work with, it is likely this will not be enough. C'est la vie.

All that said, if you don't agree, feel free, I guess to contact *Intelligence,* and explain the situation to Professor Haier. As said, we are not going to at this stage for the reasons given.

Sincerely,
John
Cc: Bryan

From: Bryan Pesta <bpesta22@cs.com>
Sent: Thursday, June 4, 2020 9:24 PM
To: j122177@hotmail.com <j122177@hotmail.com>
Subject: Fwd: Intelligence study

-----Original Message-----
From: Russell Warne <russwarne@gmail.com>
To: Bryan Pesta <bpesta22@cs.com>
Sent: Thu, Jun 4, 2020 12:01 pm
Subject: Intelligence study

Bryan,

I'm writing to make a request that is very important. I need to you remove my name from the genomics manuscript that was resubmitted to *Intelligence*.

There are two reasons for this. The first is personal. If someone found out that I was a front to gain access to the data and that I really had no hand in the creation, execution, or write-up of the study, then the effects could be drastically negative for me. Even if our actions comply with the letter, but violate the spirit of the data access agreements, the risk of negative consequences would still be too great. I'm not sure that tenure would protect me from consequences of violating the data use agreement, and I have a family to worry about. There is also the fact that my upcoming book is already causing controversy (even though it isn't even published yet), and given the current racial climate in the US, I can't afford to have any more controversies than what I'm already involved with. Primarily, this is a risk management decision on my part.

Second, there is the fact that hereditarian research will always be held to a higher standard than research supporting the environmental hypothesis. Although that is not fair, it is the reality, and it is imperative that hereditarians hold themselves to the highest scientific and ethical standards so that environmentalists cannot use irrelevant information (like an ethical lapse) to put the research and the hereditarian community into disrepute. The

short-term gain of getting a study published in a journal is not worth the potential long-term damage to hereditarianism.

I'm sure you're disappointed with this decision, but I firmly believe it is the right one. Please inform me when you have removed my contact info from the manuscript so that I can confirm that you are complying with my request.

-Russell Warne